Econometric Analysis

Ryuichi Tanaka

Ch. 7 Estimating Systems of Equations by OLS and GLS

Ex. of System Equation Model : Seemingly Unrelated Regressions (SUR)

$$\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} = \begin{pmatrix} \mathbf{x_{i1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{x_{i2}} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta}_1 \\ \boldsymbol{\beta}_2 \end{pmatrix} + \begin{pmatrix} u_{i1} \\ u_{i2} \end{pmatrix} \quad \Rightarrow \quad \mathbf{y}_i = \mathbf{x}_i \boldsymbol{\beta} + \mathbf{u}_i$$

Each regression equation looks unrelated to each other, but they are related through the correlation in the error terms $E(u_{i1}, u_{i2}) \neq 0$

Ex. Household demand functions

$$food_i = \beta_{10} + \beta_{11} foodprc_i + \beta_{12} clothprc_i + \beta_{13} income_i + u_{i1}$$
$$cloth_i = \beta_{20} + \beta_{21} foodprc_i + \beta_{22} clothprc_i + \beta_{23} income_i + u_{i2}$$

Error terms are correlated through household's budget constraint

Panel data: repeated observations for each unit

Estimation with panel data can be interpreted as an estimation of system equation model

Example 7.2 (Panel Data) : $y_t = \mathbf{x_t}\boldsymbol{\beta} + u_t,\ t = 1,2,\cdots,T$, $\mathbf{x_t} = (1, x_{1,t}, \ldots, x_{K,t}) : 1 \times K$

Why system equations?

system equations model: $y_{i1} = \mathbf{x}_{i1}\boldsymbol{\beta}_1 + u_{i1}, \ldots, y_{iG} = \mathbf{x}_{iG}\boldsymbol{\beta}_G + u_{iG}$ (G equations)

panel data model: $y_{i1} = \mathbf{x}_{i1}\boldsymbol{\beta} + u_{i1}, \ldots, y_{iT} = \mathbf{x}_{iT}\boldsymbol{\beta} + u_{iT}$ (T equations)

Ex. $saving_t = \beta_0 + \beta_1 income_t + \beta_2 age_t + \beta_3 educ_t + u_t,\ t = 1,2,\cdots,5$:

(static) linear panel data model

(quiz: which is more general? => system equations model)

Contemporaneous exogeneity and Strict exogeneity

Contemporaneous exogeneity: $E[u_t \mid \mathbf{x}_t] = 0,\ t = 1,2,\cdots,T$

Strict exogeneity: $E[u_t \mid \mathbf{x}_1, \cdots, \mathbf{x}_T] = 0,\ t = 1,2,\cdots,T$ (stronger than c-exogeneity)

Difference between contemporaneous and strict exogeneity

(1) $\mathbf{x}_t$ **contains lagged dependent variable**

Ex. pop. model : $\mathbf{x}_t = (1, y_{t-1})$: $y_t = \beta_0 + \beta_1 y_{t-1} + u_t$, $E[y_t \mid y_{t-1}, \cdots, y_0] = \beta_0 + \beta_1 y_{t-1}$

Since $E[u_t \mid \mathbf{x}_t] = E[u_t \mid y_{t-1}] = 0$, contemporaneous exogeneity is satisfied. However, since $\mathbf{x}_{t+1} = (1, y_t)$, strict exogeneity is not satisfied:

$E[u_t \mid \mathbf{x}_1, \cdots, \mathbf{x}_T] = E[u_t \mid y_0, \cdots, y_{T-1}] = E[y_t - \beta_0 - \beta_1 y_{t-1} \mid y_0, \cdots, y_{T-1}] = y_t - \beta_0 - \beta_1 y_{t-1} \neq 0$ for $t = 1,2,\cdots,T-1$.

(2) FDL model with feedback: $poverty_t = \theta_t + \delta_0 welfare_t + \delta_1 welfare_{t-1} + \delta_2 welfare_{t-2} + u_t$, $welfare_t = \eta_t + \rho_1 poverty_{t-1} + r_t$. Same problem with (1).

7.3 System OLS Estimation of a Multivariate Linear System

7.3.1 Preliminaries

Data: $\{(\mathbf{X_i}, \mathbf{y_i}) : i = 1,2,\cdots,N\}$, $\mathbf{X_i} : G \times K$, $\mathbf{y_i} : G \times 1$ (with panel data, G=T)

The population model: $\mathbf{y}_i = \mathbf{X_i}\boldsymbol{\beta} + \mathbf{u}_i$ (system equations)

Ex. $y_{i1} = \mathbf{x_{i1}}\boldsymbol{\beta_1} + u_{i1}$, $y_{i2} = \mathbf{x_{i2}}\boldsymbol{\beta_2} + u_{i2}$ => $\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} = \begin{pmatrix} \mathbf{x_{i1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{x_{i2}} \end{pmatrix} \begin{pmatrix} \boldsymbol{\beta_1} \\ \boldsymbol{\beta_2} \end{pmatrix} + \begin{pmatrix} u_{i1} \\ u_{i2} \end{pmatrix}$

With panel data: $\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} = \begin{pmatrix} \mathbf{x_{i1}} \\ \mathbf{x_{i2}} \end{pmatrix} \boldsymbol{\beta} + \begin{pmatrix} u_{i1} \\ u_{i2} \end{pmatrix}$

7.3.2 Asymptotic Properties of System OLS

<u>Assumption SOLS.1.</u> $E(\mathbf{X_i}'\mathbf{u}_i) = \mathbf{0}$

With panel data, $\mathbf{X_i}'\mathbf{u}_i = \sum_{t=1}^{T} \mathbf{x_{it}}' u_{it}$. $E(\mathbf{X_i}'\mathbf{u}_i) = \mathbf{0}$ if $E(\mathbf{x_{it}}' u_{it}) = 0$ for all $t = 1,2,\cdots,T$

ASOLS.1. is the weakest assumption we can impose in a regression framework to getconsistent estimators of beta

Weaker than contemporaneous exogeneity

<u>Assumption SOLS.2.</u> $\mathbf{A} \equiv E(\mathbf{X_i}'\mathbf{X_i})$ is nonsingular (has rank K)

ASOLS.2. is needed for identification

<u>Theorem 7.1</u> (consistency of System OLS): Under ASOLS.1. and 2, $\hat{\beta} \xrightarrow{p} \beta$ where

$$\hat{\beta} = \left(N^{-1} \sum_{i=1}^{N} \mathbf{X_i}'\mathbf{X_i} \right)^{-1} \left(N^{-1} \sum_{i=1}^{N} \mathbf{X_i}'\mathbf{y_i} \right).$$

(With panel data, SOLS estimator is equivalent to pooled OLS estimator)

<u>Theorem 7.2</u> (Asymptotic Normality of SOLS): Under ASOLS.1. and 2,

$\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} Normal(\mathbf{0}, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1})$ where $\mathbf{A} \equiv E(\mathbf{X_i}'\mathbf{X_i})$ and $\mathbf{B} \equiv E(\mathbf{X_i}'\mathbf{u_i}\mathbf{u_i}'\mathbf{X_i}) \equiv Var(\mathbf{X_i}'\mathbf{u_i})$.

<u>Estimation of Asymptotic Variance</u> : $A\mathrm{var}(\hat{\boldsymbol{\beta}}) = N^{-1}\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}$ => $\hat{\mathbf{V}} = N^{-1}\hat{\mathbf{A}}^{-1}\hat{\mathbf{B}}\hat{\mathbf{A}}^{-1}$

$\hat{\mathbf{A}} = N^{-1}\sum_{i=1}^{N} \mathbf{X_i}'\mathbf{X_i}$ **(consistent estimator of A)** $\hat{\mathbf{B}} = N^{-1}\sum_{i=1}^{N} \mathbf{X_i}'\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\mathbf{X_i}$, $\hat{\mathbf{u}}_i = \mathbf{y_i} - \mathbf{X_i}\hat{\boldsymbol{\beta}}$ **(SOLS residual)**

$\hat{\mathbf{V}}$ **can be used for t-test and robust for heteroskedasticity (including serial correlation in u)**

7.4 Consistency and Asymptotic Normality of Generalized Least Squares

GLS: In a single equation model, when $Var(u_i \mid x_i) = h(x_i)$ **is known, GLS (OLS with variables divided by** $\sqrt{h(x_i)}$ **is efficient. What happens in a system equation model?**

<u>Assumption SGLS.1.</u> $E(\mathbf{X_i} \otimes \mathbf{u}_i) = \mathbf{0}$ (each element of **u** is uncorrelated to each element of **X) Kronecker product**

Ex. $y_{i1} = \beta_{01} + \beta_{11}x_{i1} + u_{i1}$, $y_{i2} = \beta_{02} + \beta_{12}x_{i2} + u_{i2}$ <=> $\begin{pmatrix} y_{i1} \\ y_{i2} \end{pmatrix} = \begin{pmatrix} 1 & x_{i1} & 0 & 0 \\ 0 & 0 & 1 & x_{i2} \end{pmatrix} \begin{pmatrix} \beta_{01} \\ \beta_{11} \\ \beta_{02} \\ \beta_{12} \end{pmatrix} + \begin{pmatrix} u_{i1} \\ u_{i2} \end{pmatrix}$

$E(\mathbf{X_i}'\mathbf{u}_i) = \mathbf{0}$ => $E\left[\begin{pmatrix} 1 & 0 \\ x_{i1} & 0 \\ 0 & 1 \\ 0 & x_{i2} \end{pmatrix}\begin{pmatrix} u_{i1} \\ u_{i2} \end{pmatrix}\right] = E\left[\begin{pmatrix} u_{i1} \\ x_{i1}u_{i1} \\ u_{i2} \\ x_{i2}u_{i2} \end{pmatrix}\right] = \mathbf{0}$, $E(\mathbf{X_i} \otimes \mathbf{u}_i) = \mathbf{0}$ => $E\begin{bmatrix} u_{i1} & x_{i1}u_{i1} & 0 & 0 \\ u_{i2} & x_{i1}u_{i2} & 0 & 0 \\ 0 & 0 & u_{i1} & x_{i2}u_{i1} \\ 0 & 0 & u_{i2} & x_{i2}u_{i2} \end{bmatrix} = \mathbf{0}$

$E(\mathbf{u_i} \mid \mathbf{X_i}) = 0 \Longrightarrow$ ASGLS.1

Assumption SGLS.2. $\boldsymbol{\Omega} = E(\mathbf{u_i u_i}')$ is positive definite and $E(\mathbf{X_i}'\boldsymbol{\Omega}^{-1}\mathbf{X_i})$ is nonsingular.

GLS: weighted least squares with an error term with zero mean, 1 variance, and no cross-equation correlation

$\boldsymbol{\Omega}^{-1/2}\mathbf{y}_i = \boldsymbol{\Omega}^{-1/2}\mathbf{X_i}\boldsymbol{\beta} + \boldsymbol{\Omega}^{-1/2}\mathbf{u}_i$, or $\mathbf{y}_i{*} = \mathbf{X_i}{*}\boldsymbol{\beta} + \mathbf{u}_i{*}$ where $E(\mathbf{u_i}{*}\mathbf{u_i}{*}') = \mathbf{I}_G$.

$\boldsymbol{\Omega}^{1/2} = \mathbf{C}\boldsymbol{\Lambda}^{1/2}\mathbf{C}'$

$\mathbf{C}$: matrix of characteristic vectors

$\boldsymbol{\Lambda}$: diagonal matrix of characteristic roots (Greene p.973)

GLS estimator ($\boldsymbol{\beta}{*}$) is obtained by OLS of this transformed equation

Under ASGLS1 and 2, $\boldsymbol{\beta}{*}$ is <u>consistent!</u>

Moreover, $\sqrt{N}(\boldsymbol{\beta}{*} - \boldsymbol{\beta})$ is <u>asymptotic normal!</u>

7.5 Feasible GLS

For GLS, we need to know $\boldsymbol{\Omega} = E(\mathbf{u_i u_i}')$ which is usually unknown.

In FGLS, $\boldsymbol{\Omega} = E(\mathbf{u_i u_i}')$ is replaced with a consistent estimator of omega

7.5.1 Asymptotic Properties

A consistent estimator of $\boldsymbol{\Omega} = E(\mathbf{u_i u_i}')$: $\hat{\boldsymbol{\Omega}} = N^{-1}\sum_{i=1}^{N}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'$, $\hat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\hat{\boldsymbol{\beta}}}$ (***SOLS*** residuals)

Since GLS is consistent and $\underset{N\to\infty}{p\lim}\hat{\boldsymbol{\Omega}} = \boldsymbol{\Omega}$, FGLS estimator is also consistent.

Since GLS and FGLS are $\sqrt{N}$-equivalent (that is, $\sqrt{N}(\beta^{FGLS} - \beta{*}) = o_p(1)$. In other words, FGLS and GLS share the same limiting distribution), we can do asymptotic inference with FGLS as if we had GLS estimator.

<u>Theorem 7.3</u> **(Asymptotic Normality of FGLS):** Under ASGLS.1. and 2,

$\sqrt{N}(\hat{\beta} - \beta) \xrightarrow{d} Normal(\mathbf{0}, \mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1})$ where $\mathbf{A} \equiv E(\mathbf{X_i}'\boldsymbol{\Omega}^{-1}\mathbf{X_i})$ and $\mathbf{B} \equiv E(\mathbf{X_i}'\boldsymbol{\Omega}^{-1}\mathbf{u_i u_i}'\boldsymbol{\Omega}^{-1}\mathbf{X_i})$).

Estimation of Asymptotic Variance: $A\operatorname{var}(\hat{\boldsymbol{\beta}}) = N^{-1}\mathbf{A}^{-1}\mathbf{B}\mathbf{A}^{-1}$ => $\hat{\mathbf{V}} = N^{-1}\hat{\mathbf{A}}^{-1}\hat{\mathbf{B}}\hat{\mathbf{A}}^{-1}$

$\hat{\mathbf{A}} = N^{-1}\sum_{i=1}^{N}\mathbf{X_i}'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{X_i}$, $\hat{\mathbf{B}} = N^{-1}\sum_{i=1}^{N}\mathbf{X_i}'\hat{\boldsymbol{\Omega}}^{-1}\hat{\mathbf{u}}_i\hat{\mathbf{u}}_i'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{X_i}$, $\hat{\mathbf{u}}_i = \mathbf{y}_i - \mathbf{X}_i\hat{\boldsymbol{\beta}}$ **(FGLS residual)**

Notice: SOLS residual is used for the estimation of $\hat{\boldsymbol{\Omega}}$, but use FGLS residual for the estimation of asymptotic variance.

7.5.2 Asymptotic Variance of FGLS under a Standard Assumption

Under which assumption is FGLS more efficient than SOLS and other estimators?

<u>Assumption SGLS.3</u> (system homoskedasticity assumption):

$E(\mathbf{X_i}'\mathbf{\Omega}^{-1}\mathbf{u_i}\mathbf{u_i}'\mathbf{\Omega}^{-1}\mathbf{X_i}) = E(\mathbf{X_i}'\mathbf{\Omega}^{-1}\mathbf{X_i})$, where $\mathbf{\Omega} = E(\mathbf{u_i}\mathbf{u_i}')$.

<u>Theorem 7.4</u> (Usual Variance Matrix for FGLS): Under ASGLS.1, 2, and 3, the asymptotic variance of the FGLS estimator is $A\text{var}(\hat{\boldsymbol\beta}) = N^{-1}\mathbf{A}^{-1}$ where $\mathbf{A} \equiv E(\mathbf{X_i}'\mathbf{\Omega}^{-1}\mathbf{X_i})$.

$E(\mathbf{u_i}\mathbf{u_i}' \mid \mathbf{X_i}) = \mathbf{\Omega} \Leftrightarrow Var(\mathbf{u_i} \mid \mathbf{X_i}) = \mathbf{\Omega}$ **(=> ASGLS.3)**

Under SGLS.3, FGLS is more efficient than SOLS and other estimators(Problem 7.2)

<u>Theorem 7.7</u> (Large Sample Properties of Pooled OLS): Under APOLS.1. and 2, the Pooled OLS estimator is consistent and asymptotically normal. If APOLS.3. holds in addition, then $A\text{var}(\hat{\boldsymbol\beta}) = \sigma^2[E(\mathbf{X_i}'\mathbf{X_i})]^{-1}/N$, so that the appropriate estimator of $A\text{var}(\hat{\boldsymbol\beta})$ is $\hat\sigma^2(\mathbf{X}'\mathbf{X})^{-1} = \hat\sigma^2\left(\sum_{i=1}^{N}\sum_{t=1}^{T}\mathbf{x_{it}}'\mathbf{x_{it}}\right)^{-1}$ where $\hat\sigma^2$ is the usual OLS variance estimator from the pooled regression y_{it} on $\mathbf{x}_{it}$, $t = 1,2,\cdots,T$ and $i = 1,2,\cdots,N$.

Under the two assumptions, POLS is consistent. In addition, under homoskedastic assumption, asymptotic variance is given by the usual one

t-test and F-test are asymptotically valid

7.8 The Linear Panel Data Model, Revisited : $y_t = \mathbf{x_t}\boldsymbol\beta + u_t,\ t = 1,2,\cdots,T$

<u>Sufficient condition for Pooled OLS to be consistent</u>

<u>Assumption POLS.1.</u> $E(\mathbf{x}_t'u_t) = 0$ for all $t = 1,2,\cdots,T$.

<u>Assumption POLS.2.</u> $rank[\sum_{t=1}^{T} E(\mathbf{x}_t'\mathbf{x}_t)] = K$

<u>Assumption for the usual OLS inference</u>

<u>Assumption POLS.3.</u> (a) $E(u_t^2\mathbf{x}_t'\mathbf{x}_t) = \sigma^2 E(\mathbf{x}_t'\mathbf{x}_t),\ t = 1,2,\cdots,T$ where $\sigma^2 = E(u_t^2)$

(b) $E(u_t u_s \mathbf{x}_t'\mathbf{x}_s) = \mathbf{0},\ t \neq s,\ t = 1,2,\cdots,T$

Under APOLS.3., $E(\mathbf{u_i}\mathbf{u_i}') = \sigma^2\mathbf{I}_T$ **(homoskedastic)**

7.8.2 Dynamic Completeness

Dynamic Completeness of Conditional Mean: $E(y_t \mid \mathbf{x}_t, y_{t-1},\cdots,y_1,\mathbf{x}_1) = E(y_t \mid \mathbf{x}_t)$

“$\mathbf{x}_t$ contains enough information of lagged variables”

DCCM <=> $E(u_t \mid \mathbf{x}_t, u_{t-1},\cdots,u_1,\mathbf{x}_1) = 0$

=> $E(u_t u_s \mid \mathbf{x}_t,\mathbf{x}_s) = 0,\ t \neq s$

This condition $E(y_t \mid \mathbf{x}_t, y_{t-1},\cdots,y_1,\mathbf{x}_1) = E(y_t \mid \mathbf{x}_t)$ **implies APOLS3(b) and APOLS1**

Under DCCM and homoskedasticity(APOLS3(a)), we can do the usual OLS inference

7.8.5 Testing for Serial Correlation and Heteroskedasticity after Pooled OLS

Testing for Serial Correlation (in error term)

Why test serial correlation?

1. **To see if the model specification is correct (if the model has the DCCM, there must be no serial correlation)**

2. **To check if we should use variance matrix robust for serial correlation**

How to test Serial Correlation

Suppose the error term follows AR(1): $u_t = \rho_1 u_{t-1} + e_t$, where $E(e_t \mid \mathbf{x}_t, u_{t-1}, \cdots) = 0$

Plugging into the population model, $y_t = \mathbf{x_t \beta} + \rho_1 u_{t-1} + e_t$, $t = 2, \cdots, T$

$H_0 : \rho_1 = 0$ **(no serial correlation in the error term)**

Since u_{t-1} **is unobserved, replace with pooled OLS residual** $\hat{u}_{t-1}$ **and estimate** $y_t = \mathbf{x_t \beta} + \rho_1 \hat{u}_{t-1} + e_t$

Then t-test $H_0 : \rho_1 = 0$

Testing for Heteroskedasticity

Why test heteroskedasticity? Check if we should use Robust-variance matrix

How to test heteroskedasticity

Assume $E(u_t \mid \mathbf{x}_t) = 0$ **(stronger than APOLS1). The the null hypothesis is** $H_0 : E(u_t^2 \mid \mathbf{x}_t) = \sigma^2$**. Under the null,** u_{it}^2 **has no correlation with any non-constant function (**$\mathbf{h}_{it}$**) of** $\mathbf{x}_{it}$

Construct squared pooled OLS residual $\hat{u}_{it}^2$**. Regress it on 1 and** $\mathbf{h}_{it}$ **and obtain** R_c^2**. Under the** H_0**,** NTR_c^2 **follows asymptotically** χ_Q^2 **(where Q is the dimension of h)**

7.8.6 FGLS Estimation under Strict Exogeneity

Occasionally we assume strict exogeneity.

However, it is not satisfied in many cases (ex. the model with $\mathbf{x}_t$ **containing lagged dependent variable, or Finite Distribution Lag (FDL) model with feedback)**

We need to be careful when we assume strict exogeneity

PS 7

I. 7.2

II. Reproduce the table 7.1 in Example 7.3 (hint: use the stata command sureg

III. 7.8

IV. Reproduce the estimation results in Example 7.7 with Stata (hand in both the output and the DO files)

V. 7.9